

Conference Abstract

Anatomy Knowledge Graphs: Toward FAIR morphological data

Lars Vogt [‡]

[‡] Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany

Corresponding author: Lars Vogt (lars.m.vogt@gmail.com)

Received: 12 Jun 2019 | Published: 19 Jun 2019

Citation: Vogt L (2019) Anatomy Knowledge Graphs: Toward FAIR morphological data. Biodiversity Information Science and Standards 3: e37203. <https://doi.org/10.3897/biss.3.37203>

Abstract

Most morphological data are still published as unstructured texts. This has far-reaching consequences for the Findability, Accessibility, Interoperability and Reusability of morphological data and thus for their FAIRness (Wilkinson et al. 2016). The lack of FAIR morphological data significantly affects their general usability within the life sciences. With the advent of the Semantic Web and an increasing amount of publicly accessible anatomy ontologies, technically feasible solutions to this linguistic problem of morphology have become available. After a brief introduction to the distinction between instance anatomy (anatomical data pertaining to instances, i.e. individuals) and canonical anatomy (generalized anatomical knowledge), between assertional statements (i.e., ABox expressions, statements about instances) and universal statements (i.e., TBox expressions, statements about kinds or classes), and between knowledge bases and ontologies, I compare two different approaches of representing anatomy through Resource Description Framework (RDF) based graphs (see Fig. 1): i) *Semantic Phenotypes*, which are class-based graphs that document anatomy purely as TBox expressions and ii) *Semantic Instance Anatomy Knowledge Graphs* (short: *Anatomy Knowledge Graphs*), which are instance-based graphs that document anatomy as a combination of ABox and TBox expressions, thereby limiting the portion of TBox expressions to a minimum. While both approaches provide solutions to the linguistic problem of morphology, they differ not only conceptually but also in their technical details. Because *Anatomy Knowledge Graphs* include ABox expressions, they allow the identification of any individually described part or property within the description through the part's or property's particular URI. This detail

allows fragmentation of a morphological description into meaningful description fragments and reuse of these fragments in other contexts. Because *Anatomy Knowledge Graphs* limit TBox expressions to a minimum, they also perform better in many queries as querying ABox expressions is computationally less difficult than querying TBox expressions. These technical differences between the two approaches have substantial practical consequences that significantly affect the FAIRness of morphological data. Compared to *Semantic Phenotypes*, morphological data in the form of *Anatomy Knowledge Graphs* are overall more FAIR.

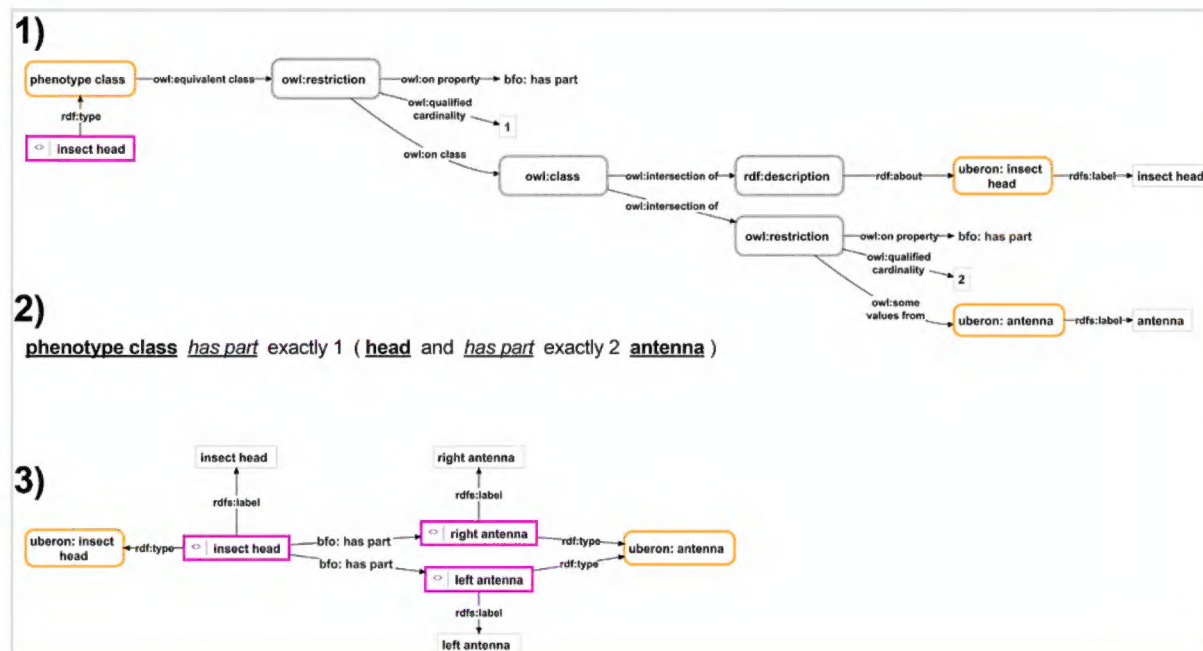


Figure 1.

Phenotype description of an insect head with two antennae in form of a *Semantic Phenotype* in the form of an RDF-based graph or in Manchester syntax, and an *Anatomy Knowledge Graph*.

1) The *Semantic Phenotype* description of an insect head that has two antennae as its parts in the form of an RDF-based graph. It consists of an instance (purple-bordered box) that instantiates the phenotype class that contains the actual description of the phenotype (yellow-bordered box) in form of a class axiom consisting of anonymous property restrictions and class descriptions (grey-bordered boxes). The class axiom characterizes all instances of the class to consist of exactly one instance of insect head (id UBERON:6000004) that has as its parts exactly two instances of antenna (id UBERON:0000972). **2)** An alternative format of the same *Semantic Phenotype* representing the class axioms from the 'phenotype class' from (1) expressed in OWL Manchester Syntax (ontology classes shown with their label in bold and underlined, ontology properties with their label in italics and underlined, 'and' being used in the sense of intersection of two mathematical sets and 'exactly' as a cardinality specification). **3)** The *Anatomy Knowledge Graph* description of an insect head that has two antennae as its parts. It consists of the instance of insect head (id UBERON:6000004) and two instances of antenna (id UBERON:0000972), which relate to the head as its parts. Labels (in light-grey-bordered boxes) indicate how the different instances should be represented in a human-readable format. For reasons of clarity, resources are not represented with their URIs but with labels. *purple-bordered box = ontology instance; yellow-bordered box with rounded corners = ontology class; grey-bordered box with rounded corners = anonymous class; light-grey-bordered box = literal or numerical value; labeled arrow = property.*

Keywords

Anatomy Knowledge Graph, semantic phenotype, FAIR data principle, anatomy, morphological description, semantic graph, knowledge graph, ontology

Presenting author

Lars Vogt

Presented at

Biodiversity_Next 2019

Acknowledgements

I thank Peter Grobe, Roman Baum, Christian Köhler, Sandra Meid, Björn Quast, Stefan Richter, Christian Wirkner, Michael Ohl, Stefan Graf, Jan Decher, and Christian Montermann for discussing the here presented ideas. I thank István Mikó, Matt Yoder, and Torben Göpel for discussing some of the data schemes for representing morphological structures.

Funding program

VO 1244/8-1

Grant title

eScience-Compliant Standards for Morphology

Hosting institution

Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany

References

- Wilkinson M, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J, da Silva Santos LB, Bourne P, Bouwman J, Brookes A, Clark T, Crosas M, Dillo I,

Dumon O, Edmunds S, Evelo C, Finkers R, Gonzalez-Beltran A, Gray AG, Groth P, Goble C, Grethe J, Heringa J, 't Hoen PC, Hooft R, Kuhn T, Kok R, Kok J, Lusher S, Martone M, Mons A, Packer A, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S, Schultes E, Sengstag T, Slater T, Strawn G, Swertz M, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016) The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3: 160018. <https://doi.org/10.1038/sdata.2016.18>